

## The Information Sciences Series

*Edited by* ROBERT M. HAYES, *University of California at Los Angeles,*  
*and* JOSEPH BECKER, *Interuniversity Communications Council*  
*(EDUCOM); Consultants:* CHARLES P. BOURNE, *Director, Advanced*  
*Information Systems Division Programming Services, Inc., and* HAROLD  
BORKO, *System Development Corporation.*



This series covers a broad spectrum of topics dealing with the storage and use of digital and graphic information, including: information system theory and design . . . man-machine relationships . . . language data processing . . . artificial intelligence . . . mechanization of library processes . . . non-numerical application of digital computers . . . automatic publishing . . . and information display.

ANNOUNCING

*the latest volume in*  
*this continuing series —*

## The Analysis of Information Systems

**A Programmer's Introduction to  
Information Retrieval**

**CHARLES T. MEADOW**

Federal Systems Division

International Business Machines Corporation

# **The Analysis of Information Systems A Programmer's Introduction to Information Retrieval**

By CHARLES T. MEADOW, *Federal Systems Division, IBM Corporation.*

Information retrieval languages, the organization of records and files, and the processing of files all have important effects on the performance and economics of retrieval systems. This book provides a careful treatment of the problems associated with these three aspects of using a computer to store and retrieve information.

Approaching information retrieval as a communication activity between system user, library, and author, the book deals first with the available languages — including the language of the document, the library's description of that document in diverse classification and indexing languages, and the use of language to query information files.

Types of record and file structure are then considered with regard to their ability to represent the languages and their advantages and disadvantages in file searching. Finally, the book presents techniques for performing the standard file processing functions of sorting, merging, and file maintenance.

1967.

301 pages.

\$11.50.

## **Contents**

### **Part I — INFORMATION RETRIEVAL AND COMMUNICATION**

#### **INTRODUCTION**

Information Retrieval as a Communication Process  
The Information Retrieval Process

#### **THE LANGUAGES OF INFORMATION RETRIEVAL**

Introduction  
Attributes of Index Languages  
A Survey of Index Languages  
The Logic of Indexing Languages  
A Comparative Analysis of Index Languages  
Dictionaries

#### **THE CREATION OF INDEX RECORDS**

Information  
Principles of Index Construction  
Use of Document Structure in Indexing  
Statistical Selection of Index Terms

### **RETRIEVAL OF INDEX RECORDS**

Introduction  
Principles of Retrieval  
Retrieval Planning  
Retrieval System Performance  
Use of Retrieval Information

### **Part II — THE ORGANIZATION OF INFORMATION**

#### **THE ORGANIZATION OF RECORDS**

Introduction  
Principles of Record Organization  
The Progression of Record Structures  
Summary

#### **THE ORGANIZATION OF FILES**

Introduction  
File Processing and Searching  
The Role of Storage Media in File Organization  
Principles of File Organization  
Methods of File Organization

#### **THE STORAGE AND MEASUREMENT OF STRUCTURAL INFORMATION**

Introduction  
Concepts of Information Theory  
Representation of File Structure Information  
Measurement of File Structure Information

#### **THE ORGANIZATION OF FILE SETS**

Introduction  
Principles of File Set Organization  
Fixed Sequence File Searching  
The Inverted Index File  
Variable Sequence Searching

### **Part III — PROCESSING OF FILES AND FILE SETS**

#### **BASIC FILE PROCESSES**

Introduction  
File Sequence  
Merging  
Sorting  
Searching  
Batch Processing

(Continued overleaf)

## FILE MAINTENANCE

Introduction  
File Maintenance Principles  
Examples of File Maintenance Processing  
Summary

## The Information Sciences Series

### NETWORK-BASED MANAGEMENT SYSTEMS (PERT/CPM)

By RUSSELL D. ARCHIBALD, *Informatics, Inc.*, and RICHARD VILLORIA, *Houston-Fearless Corporation*.

1967. 508 pages. \$16.00.

### INFORMATION STORAGE AND RETRIEVAL

By JOSEPH BECKER, *Interuniversity Communications Council (EDUCOM)*, and ROBERT M. HAYES, *University of California, Los Angeles*.

1963. 448 pages. \$11.95.

### AUTOMATED LANGUAGE PROCESSING

By HAROLD BORKO, *System Development Corporation*.

1967. Approx. 480 pages. Prob. \$12.95.

### METHODS OF INFORMATION HANDLING

By CHARLES P. BOURNE, *Programming Services Inc.*

1963. 241 pages. \$12.95.

### NATIONAL DOCUMENT-HANDLING SYSTEMS FOR SCIENCE AND TECHNOLOGY

By LAUNOR F. CARTER, *System Development Corporation*.

1967. 344 pages. \$9.95.

### MATHEMATICS OF INFORMATION SYSTEMS

By ROBERT M. HAYES, *University of California, Los Angeles*.

*In preparation.*

### THE ANALYSIS OF INFORMATION SYSTEMS:

A Programmer's Introduction to  
Information Retrieval

By CHARLES T. MEADOW, *IBM Corporation*.

1967. 301 pages. \$11.50.



**John Wiley & Sons, Inc.**

**605 Third Avenue**

**New York, N. Y. 10016**



Wiley presents the most recent

# information

*for all  
progress-oriented  
users...  
designers...  
and students of  
information  
systems & services*

# on information

*from COSATI...American Documentation Institute...System  
Development Corporation...and National Science Foundation*

*The results of the COSATI-sponsored study on national information systems relating to scientific and technical documents—presented in a clear and well-organized manner.*

## **National Document-Handling Systems for Science and Technology**

By LAUNOR F. CARTER, GORDON CANTLEY, JOHN T. ROWELL, LOUISE SCHULTZ, HERBERT R. SEIDEN, EVERETT WALLACE, RICHARD WATSON, and RONALD E. WYLLYS, *all of the System Development Corporation, Santa Monica, California.*

"This book has grown out of a study undertaken for the Committee on Scientific and Technical Information (COSATI) Task Group on National System(s) for Scientific and Technical Information in support of their examination of national document and information systems. The emphasis of the study as stated by COSATI is the following:

" '1. Initial and primary priority will be placed on national systems relating to scientific and technical documents, their handling and the management of such documents. . . .

" '2. Secondary attention will be given to development of programs which can be undertaken with Government support for identifying, analyzing and giving a structure to the total flow of scientific and technical information in the United States.' "

*From the Preface.*

The System Development Corporation group that participated in this study collected information from reports of former study groups, published studies of user needs for scientific and technical information, and from observations of the scientific and technical information system as it is presently functioning in public and private institutions.

*National Document-Handling Systems for Science and Technology* contains all the ideas presented in the original COSATI report—expanded and clarified by the people initially responsible for them—in a format and order that has been extensively revised to make the book more informative and readable.

☞ This is the kind of information you'll find in this book

A complete study of information-handling organizations and systems as they operate in the United States today

A survey of the degree to which the output of the scientific and technical community is decreased by the deficiencies in these organizations and systems

A wealth of useful details about available and economically feasible systems that would correct these deficiencies

Five action plans for the future—drawn up by the COSATI experts—for use by governmental and private institutions

#### **CONTENTS:**

Introduction

#### **Description of the Present System**

The Present System: Document-Handling Institutions. The Present System: Document Flow. Document Users.

#### **The Problem and Basic Propositions**

The National Document-Handling Problem. Basic Propositions and System Requirements.

#### **Alternative Approaches to the Problem**

Selected Plans for National Document Systems. The Capping Agency and Responsible Agent Concepts. The Concept of New Operating Agencies. The Concept of Strengthening the Present System.

#### **An Evaluation and Prognosis**

An Evaluation of Alternatives. Present Status and Future Programs.

#### **Appendices**

Advanced Information Systems, Equipment and Software. Statistical Review. Review of Federal Legislation and Executive Orders.

Index.

1967.

344 pages.

\$9.95.

*A publication in the Wiley Information Science series.*



A complete annual review—sponsored by the American Documentation Institute.

## Annual Review of Information Science and Technology

Edited by CARLOS A. CUADRA, System Development Corporation.

In 1964, the National Science Foundation joined the System Development Corporation in lending major support to the initiation of an annual review series to be concerned with information science and technology. The American Documentation Institute accepted sponsorship of the reviews and appointed an editor and an advisory committee early in 1965.

The series is intended to be of tangible benefit to individuals, public and private organizations, universities, and government agencies. Although the field cuts across numerous disciplines, the people involved in it all share a concern with the generation, transformation, communication, storage, retrieval, and use of information. This volume and the ones which will follow it, are planned as constructive reviews of topics of interest to users, designers and students of information systems and services. By providing a perspective on the information field, such reviews not only point out gaps and duplications but also serve as sources of new ideas. The *Annual Review* will attempt to reflect current interests, and, by including information on the field from a number of related disciplines, to broaden and deepen them.

The present volume covers, for the most part, literature that appeared in 1965. Some of the earlier literature is also reviewed because this volume, as the first in the series, had no prior context to serve as a frame of reference. Volume II, which will appear in the fall of 1967, will cover 1966 literature.

### CONTENTS and AUTHORS

Introduction to the ADI Annual Review—Carlos A. Cuadra. Professional Aspects of Information Science and Technology—Robert S. Taylor. Information Needs and Uses in Science and Technology—Herbert Menzel. Content Analysis, Specification and Control—Phyllis Baxendale. File Organization and Search Techniques—W. Douglas Climenson. Automated Language Processing—Robert F. Simmons. Evaluation of Indexing Systems—Charles P. Bourne. New Hardware Developments—Annual Review Staff. Man-Machine Communication—Ruth M. Davis. Information System Applications—Jordan J. Baruch. Library Automation—Donald V. Black and Earl Farley. Information Centers and Services—G. S. Simpson, Jr. and Carolyn Flanagan. National Information Issues and Trends—John Sherrod. Index—Pauline Atherton and Stella Keenan.

1966. 389 pages. \$12.50.

ADDITIONAL COPIES

Typeset and indexed by computer—a book that shows how to use computers to process natural language in order to facilitate the communication, production, translation, storage and retrieval of information.

## **Automated Language Processing**

### **The State of the Art**

*Edited by* HAROLD BORKO, *System Development Corporation.*

In the light of simultaneously increasing knowledge and specialization—especially in science and technology—the increased efficiency of information flow has come to be of primary importance. In the age of automation this can be successfully accomplished only by automated means—language processing by computer. The objective of this book is—by presenting basic information on the procedures and purposes of processing natural language data by computer—to identify and define the problems which remain to be solved if the system is to be completely adequate.

In the first section, functions basic to language processing are related to computer processing systems, and the applications of mathematical models to language and linguistics are examined. The applications of statistical processing techniques to indexing, abstracting and analysis of style are considered in the second part. Finally, syntactic analysis techniques and their use in computer question answering, translation, and information retrieval are discussed.

The book itself is an experiment in language data processing: it has been typeset and indexed by computer.

Each chapter has been written expressly for this book by a recognized leader in his field.

#### **CONTENTS and AUTHORS:**

##### **Language Data Processing**

Introduction—*Harold Borko*. Language and the Computer—*Louise Shultz*. Mathematical Models of Language—*H. P. Edmundson*.

##### **Statistical Analysis**

Indexing and Classification—*Harold Borko*. Extracting and Abstracting by Computer—*Ronald E. Wyllys*. Stylistic Analysis—*Sally and Walter Sedelow*.

##### **Syntactic Analysis**

Theories in Computer Implementations—*D. G. Bobrow*. Answering English Questions by Computer—*R. F. Simmons*. Translating Languages—*E. D. Pendergraft*. Designing Artificial Languages for Information Storage and Retrieval—*C. Kellogg*.  
1967. Approx. 480 pages. Prob. \$12.95.

Order from your bookseller—or order direct from Wiley by asking for F08950 Borko: Language.

*A publication in the Wiley Information Sciences series*



"... the most up-to-date review of mechanized information retrieval..."  
—JOHN MARKUS, McGraw Hill Book Company, in *Proceedings of the IEEE*.

Now available in a new edition

## Textbook on Mechanized Information Retrieval

By ALLEN KENT, Knowledge Availability Systems Center and the University of Pittsburgh.

### CONTENTS:

Introduction. Physical Tools—Introduction. Physical Tools—Input Operations. Physical Tools—Output Operations. Principles of Analysis. Principles of Searching. Manipulation of Searching Devices. Words, Language, and Meaning in Retrieval Systems. Codes and Notations. Systems Design Criteria. Data Processing and the Library. Research on Scientific Documentation and Information Retrieval. Appendix. Author Index. Subject Index.

1966. 371 pages. \$10.95.

Volume 3 in the INTERSCIENCE Library Science and Documentation series.

## The Analysis of Information Systems

### A Programmer's Introduction to Information Retrieval

By CHARLES T. MEADOW, International Business Machines Corporation.

Deals with the problems and methods of using a computer to store, search and retrieve information—for those involved in the analysis, design or use of information systems.

### CONTENTS:

#### Information Retrieval and Communication:

Introduction. The Languages of Information Retrieval. The Creation of Index Records. Retrieval of Index Records.

#### The Organization of Information:

The Organization of Records. The Organization of Files. The Storage and Measurement of Structural Information. The Organization of File Sets.

#### Processing of Files and File Sets:

Basic File Processes. File Maintenance. 1967. 301 pages. \$11.50.

A publication in the Wiley Information Science series.

## Information Theory

By ROBERT ASH, University of Illinois.

"My opinion is that it is the best textbook so far available on this subject at this level. I am pleased that he manages to be mathematically correct as well as interesting." Professor Leo Breiman, University of California.

### CONTENTS:

A Measure of Information. Noiseless Coding. The Discrete Memoryless Channel. Error Correcting Codes. Further Theory of Error Correcting Codes. Information Sources. Channels with Memory. Continuous Channels. Appendices: Compact and Symmetric Operators on  $L_2[a,b]$ . Integral Operators. The Karhunen-Loève Theorem. Further Results Concerning Integral Operators Determined by a Covariance Function. 1965. 339 pages. \$13.50.

INTERSCIENCE: Tracts in Pure and Applied Mathematics, Volume 13.

"... the most useful book in English on the subject..."

College and Research Libraries

## Information Storage and Retrieval:

### Tools, Elements, Theories

By JOSEPH BECKER, Data Processing and Library Consultant, and ROBERT M. HAYES, University of California, Los Angeles.

### CONTENTS:

TOOLS: Introduction. The Librarian and Recorded Knowledge. The Documentalist and the Development of New Techniques. The Information Framework and the User. Printed Data and the Creation of a Machine Language. Analysis, Logical Processing and the Computer. Indexes, Documents, and the Storage of Data.

ELEMENTS: Interdisciplinary Character of Information Systems. Elements of Usage. Elements of Organization. Elements of Equipment. Parameters and Implementation.

THEORIES: Role of a Theory. Theories of File Organization. Theories of System Design. 1963. 448 pages. \$11.95.

A publication in the Wiley Information Science Series.



JOHN WILEY & SONS, Inc.  
605 Third Avenue, New York, N.Y. 10016

Please send the books I have checked on ten days' approval. I agree to pay for the books, plus a few cents for shipping charges, or return them postpaid, in good condition, within ten days of receipt. (Wiley pays postage on orders accompanied by remittance.)

☐ F 13731 Carter: National Document  
☐ I 18919 Cuadra: Information, Vol. 1  
☐ F 08950 Borko: Language  
☐ I 46935 Kent: Textbook, 2nd Ed.  
☐ F 59000 Meadow: Information

☐ I 03445 Ash: Information  
☐ F 06129 Becker: Information  
☐ F 09150 Bourne: Information  
☐ I 83310 Strauss: Libraries

Send to:

Firm \_\_\_\_\_

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Z.C. \_\_\_\_\_

☐ Payment Enclosed

# Methods of Information Handling

By CHARLES P. BOURNE, *Advanced Information Systems Division, Programming Services, Incorporated.*

"Bourne aims at providing 'an aid and reference work for those people who are interested in the design of information systems.' The aim is amply fulfilled, and the work is highly recommended. . . . It shows a broad and eclectic knowledge of the literature."  
*American Documentation.*

## CONTENTS:

The Nature of the Problem. Classification and Indexing: The Organization of Information. Coding: The Indexing Shorthand. Machine-Language Representation. Manual Card Systems. Punched Card Systems. Computer Systems. Other Paper-Tape and Magnetic Media Equipment. Microfilm and Image Handling Equipment.

1963. 241 pages. \$12.95

*A publication in the Wiley Information Science Series.*

*Also available*

# Scientific and Technical Libraries

## *Their Organization and Administration*

By LUCILLE J. STRAUSS, IRENE M. STRIBY, and ALBERT L. BROWN.

"... the kind of reference the wise technical library administrator will keep on the shelf as required reading ..."  
*Special Libraries*

1964. 398 pages. \$8.50.

*A volume in the INTERSCIENCE Library Science and Documentation series.*



## Business Reply Card

FIRST CLASS PERMIT NO. 32664, NEW YORK, N. Y.

**JOHN WILEY & SONS, Inc.**  
605 Third Avenue  
New York, N.Y. 10016

*Attention: I. Rosen*





**Editorial Advisory  
Board:**

JOHN MAKENZIE CORY

*Chief, Circulating Department, The  
New York Public Library*

LUCILE L. KECK

*Librarian, Joint Reference Library*

JAMES W. PERRY

*Numerical Analysis Laboratory, Uni-  
versity of Arizona*

IRENE M. STRIEBY

*Library Consultant, formerly Chief  
Librarian, Eli Lilly and Company*

MAURICE F. TAUBER

*School of Library Science, Columbia  
University*

EILEEN THORNTON

*Oberlin College Library*

# **Library Science and Documentation**

## **A Series of Texts and Monographs**

*General Editor: JESSE H. SHERA, Dean, School of Library Science, Western  
Reserve University.*

A new edition of Volume 3

### **Kent: TEXTBOOK ON MECHANIZED INFORMATION RETRIEVAL**

is now available



# Textbook on Mechanized Information Retrieval

## Second Edition

By ALLEN KENT, Director, Knowledge Availability Systems Center and Professor of Library Science and Education at the University of Pittsburgh.

Clarification, up-dating, and augmentation of material distinguish the second edition of *Textbook on Mechanized Information Retrieval*. The first edition was designed as a thorough introduction to the developing IR field; the second edition retains the orientation of its predecessor and offers these new features:

- physical tools presented in three chapters instead of one. This division facilitates fuller treatment of necessary background material and input and output operations.
- significant additions to chapters 1, 2, and 5—especially on abstracting and processing of full texts.
- new diagrams illustrating search strategy principles in chapter 6.
- freshly-conceived discussion of direct, indirect, and superimposed coding in chapter 9. This exposition aids the beginner in learning how to choose among these approaches when designing a coding system.
- chapter 10 ("Systems Design Criteria") and chapter 11 ("Data Processing and the Library") are entirely new.

*Textbook on Mechanized Information Retrieval* in its second edition is still the only book that *teaches* the fundamentals of the IR field without requiring a special mathematical background. (The reader receives an introduction to Boolean algebra via the discussions of search strategies.) Finally—this book is unique by virtue of the fact that it considers the information retrieval field from the point of view of unit operations of IR systems, *i.e.*, acquisition, analysis, vocabulary control, recording of results of analysis in searchable media, storage of source materials, etc. Both the human and the machine aspects of each unit operation are presented.

## Chapter Headings

Introduction  
Physical Tools—Introduction  
Physical Tools—Input Operations  
Physical Tools—Output Operations  
Principles of Analysis  
Principles of Searching  
Manipulation of Searching Devices  
Words, Language, and Meaning in Retrieval Systems  
Codes and Notations  
Systems Design Criteria

1966. 371 pages. \$10.95.

Send for your copy today.

*Other volumes in the series*

## **SCIENTIFIC AND TECHNICAL LIBRARIES Their Organization and Administration**

by LUCILLE J. STRAUSS, *The Pennsylvania State University*;  
IRENE M. STRIEBY, *Consultant in the Organization of  
Special Libraries and Company Archives*; and  
ALBERTA L. BROWN, *Western Michigan University*.

A unique and invaluable handbook that describes the principles, practices, and tools required for the maintenance of scientific and technical libraries. Special attention is given to the problems and needs of small- and medium-sized operations. (*Library Science and Documentation—Volume 4*).

1964. 398 pages. \$8.50.

## **PUNCH-CARD METHODS IN RESEARCH DOCUMENTATION With Special Reference to Biology**

By MARTIN SCHEELE, *Hydrobiologische Anstalt der Max-Planck  
Gesellschaft*; translated from the German by EDWIN HOLMSTROM.

Synthesizing principles from various punch-card methods, this book develops a series of general rules governing punch-card applications to research and documentation. Although the role of punch-cards in biological research is stressed, the author also discusses their role in many other fields, such as meteorology, aerodynamics, and chemistry. (*Library Science and Documentation—Volume 2*).

1962. 282 pages. \$9.50.

## **TOOLS FOR MACHINE LITERATURE SEARCHING Semantic Code Dictionary•Equipment•Procedures**

By J. W. PERRY, *University of Arizona*; and ALLEN KENT,  
*University of Pittsburgh*; *Semantic Code Dictionary Prepared under  
the General Editorship of JOHN L. MELTON, John Carroll University*.

A comprehensive guide to the use of modern automatic equipment in literature searching. (*Library Science and Documentation—Volume 1*).

1958. 992 pages. \$27.50.

## **INTERSCIENCE PUBLISHERS**

*a division of*

JOHN WILEY & SONS, Inc.,

605 Third Avenue, New York, N. Y. 10016

10083

Postage  
Will be Paid  
by  
Addressee

No  
Postage Stamp  
Necessary  
If Mailed in the  
United States

BUSINESS REPLY MAIL  
First Class Permit No. 32664, New York, N. Y.

*John Wiley & Sons, Inc.*

*605 Third Avenue*

*New York, N. Y. 10016*



PLEASE SEND the following books on 10 days' approval. I agree to pay for the books, plus a few cents' postage and handling, or return them postpaid, in good condition, within 10 days of receipt. (Wiley pays postage on orders accompanied by remittance.)



Print the names of the books you want:

(Author)

(Title)

---

---

---

---

---

---

---

---

**BILL TO:**

Firm

Name

Address

City

Zip

State

**SHIP TO** (*If same as BILL TO, write SAME*)

Firm

Name

Address

City

Zip

State

Professional books are tax deductible.

☐ Payment enclosed